Crystal structures of the Gon7/Pcc1 and Bud32/Cgi121 complexes provide a model for the complete yeast KEOPS complex.

Wenhua Zhang, Bruno Collinet, Marc Graille, Marie-Claire Daugeron, Noureddine Lazar, Domenico Libri, Dominique Durand, Herman van Tilbeurgh

3358–3372 *Nucleic Acids Research, 2015, Vol. 43, No. 6* *Published online 03 March 2015*
doi: 10.1093/nar/gkv155

Crystal structures of the Gon7/Pcc1 and Bud32/Cgi121 complexes provide a model for the complete yeast KEOPS complex

Wenhua Zhang[1], Bruno Collinet[1,2], Marc Graille[1], Marie-Claire Daugeron[3], Noureddine Lazar[1], Domenico Libri[4], Dominique Durand[1] and Herman van Tilbeurgh[1,*]

[1]Institut de Biologie Intégrative de la Cellule, UMR 9198, CNRS, Université de Paris Sud XI, Bâtiment 430, 91405 Orsay, France, [2]Sorbonne Universités, UPMC Univ Paris 06, UFR 927, Sciences de la vie, F-75005, Paris, France, [3]Domenico Libri 33 Institut Jacques Monod, CNRS, UMR 7592, Université de Paris Diderot, Sorbonne Paris Cité, 75205 Paris, France and [4]Institut Jacques Monod, CNRS, UMR 7592, Université de Paris Diderot, Sorbonne Paris Cité, 75205 Paris, France

Received January 13, 2015; Revised February 16, 2015; Accepted February 18, 2015

ABSTRACT

The yeast KEOPS protein complex comprising Kae1, Bud32, Cgi121, Pcc1 and Gon7 is responsible for the essential tRNA threonylcarbamoyladenosine (t^6A) modification. Deletion of genes coding for the KEOPS subunits also affects telomere elongation and transcriptional regulation. In the present work, the crystal structure of Bud32/Cgi121 in complex with ADP revealed that ADP is bound in the catalytic site of Bud32 in a canonical manner characteristic of Protein Kinase A (PKA) family proteins. We found that Gon7 forms a stable heterodimer with Pcc1 and report the crystal structure of the Pcc1-Gon7 heterodimer. Gon7 interacts with the same Pcc1 region engaged in the archaeal Pcc1 homodimer. We further show that yeast KEOPS, unlike its archaeal counterpart, exists as a heteropentamer in which Gon7, Pcc1, Kae1, Bud32 and Cgi121 also adopt a linear arrangement. We constructed a model of yeast KEOPS that provides structural insight into the role of Gon7. The model also revealed the presence of a highly positively charged crater surrounding the entrance of Kae1 that likely binds tRNA.

INTRODUCTION

The N^6-threonylcarbamoyladenosine (t^6A) modification is universally present at position 37 of tRNAs that recognize ANN-codons, with N being any nucleotide (1,2). The t^6A modification enhances the codon–anti-codon interaction, is required for recognition of the AUG start codon and is important for maintaining the translational fidelity (3,4). Abolition of the t^6A modification leads to increased frame shift events and errors in start codon recognition (4). Impaired t^6A modification therefore affects the translation of many proteins and hence might indirectly impact multiple cellular processes (5,6). Comparative genomics and biochemical experiments identified the enzymes involved and provided insight into the reaction mechanism. The t^6A biosynthesis reaction proceeds in two steps. The universal Sua5 enzyme first synthesizes an unstable threonylcarbamoyladenylate (TCA) intermediate from adenosine triphosphate (ATP), threonine and bicarbonate. In the second step, the threonylcarbamoyl-moiety of TCA is transferred onto A_{37} of substrate tRNA. The TCA transfer is catalyzed in archaea and eukaryotes by the multi-subunit complex KEOPS (6–8), which in yeast is composed of five proteins: Gon7, Pcc1, Kae1, Bud32 and Cgi121 (19). There exists ample evidence that the TCA transfer takes place on Kae1 but the function of the other KEOPS subunits remains largely unclear. Qri7 is the yeast mitochondrial paralog of Kae1 and is able to complement Kae1 in the t^6A biosynthesis (4). A recent study showed that Sua5 and Qri7 constitute the minimal system for the *in vitro* t^6A biosynthesis (8). In bacteria, the *in vitro* t^6A biosynthesis requires YrdC (Sua5 ortholog), YgjD (TsaD, Kae1 ortholog) and two bacterial-specific proteins YeaZ (TsaB) and YjeE (TsaE)(7,10,11).

The crystal structures of archaeal Kae1 (12), mitochondrial Qri7 (8) and bacterial YgjD (13) showed that they belong to the ASKHA (acetate and sugar kinases/Hsp70/actin) superfamily (14). A non-heme Fe(III) sits in the active site of *Pyrococcus abyssi* Kae1, liganded by two histidines and one aspartate, which form

*To whom correspondence should be addressed. Tel: +33 1 69 15 31 55; Email: herman.van-tilbeurgh@u-psud.fr
Correspondence may also be addressed to Bruno Collinet. Tel: +33 1 69 15 79 68; Email: bruno.collinet@u-psud.fr
Present address: Marc Graille, Laboratoire de Biochimie, CNRS, UMR 7654, Ecole Polytechnique, F-91128 Palaiseau Cedex, France.

a conserved metal-binding motif in Kae1/Qri7/YgjD (4). The metal directly contacts the γ- and β-phosphates of the AMPPNP in archaeal Kae1 (15,16) and the ATPγS in *Salmonella typhimurium* and *Escherichia coli* YgjD (11,13). The crystal structures of archaeal Pcc1, Kae1, Bud32 and Cgi121 as well as those of a few KEOPS subcomplexes have been determined (16,17). The ensemble of structures of subcomplexes allowed reconstructing a structural model of KEOPS from archaea (17). KEOPS displays a linear rearrangement of subunits, with Pcc1 at one end, Kae1 and Bud32 in the middle and Cgi121 at the other end. The N-terminal domain of Kae1 binds to Pcc1 and its C-terminal domain protrudes into the two lobes of the Bud32 kinase protein. Bud32 is structurally related to the small-size atypical RIO kinases that lack the long activation loop and are involved in ribosome biogenesis (18). Cgi121 binds to the N-terminal lobe of Bud32 and has presently no structural analogs. Archaeal Pcc1 functions as a dimerization module that further mediates the formation of a dimer of heterotetrameric KEOPS, exhibiting a V-shape molecular architecture in which Pcc1 dimer is situated at the base connecting the two arms (17). Finally, the fifth component of KEOPS complex from yeast is Gon7 for which no orthologs have been found outside the fungal lineage. Nothing is known about its specific function, though it is required for t^6A biosynthesis and telomere length maintenance (6,19). Gon7 gene expression can be induced during osmotic stress (20) and the Gon7-knockout yeast strain shows a defect in mannosyl phosphorylation of mannoprotein-linked oligosaccharides (21).

Telomeres are the physical ends of linear chromosomes and they shorten along with genome replication during cell division. Telomere capping in yeast is carried out by a protein complex containing Cdc13, Stn1 and Ten1 (22,23). In a search for new *Saccharomyces cerevisiae* genes involved in capping, KEOPS was shown to promote both telomere uncapping and elongation (17,19). In a parallel-unrelated study, Kisseleva-Romanova *et al.* identified the same complex and documented its role in transcriptional regulation of GAL genes in yeast (24). Interaction of the KEOPS complex with chromatin suggested that it could also function as a transcription factor (24). The functional importance of the integrity of the yeast KEOPS complex was tested *in vivo*. Mutations that lead to disruption of the Kae1/Bud32 and Pcc1/Kae1 complexes had pleiotropic phenotypes including slow growth, telomere shortening, defective transcriptional regulation, similar to effects caused by deletion of either KAE1 gene or BUD32 gene (5,16,17). Disruption of the Cgi121/Bud32 also caused slow growth and short telomeres but the effect was milder than that for Kae1/Bud32 (17). It is for the moment not known whether the pleiotropic effects of the KEOPS deletion mutations are a consequence of the defects in t^6A metabolism.

The mechanism of tRNA t^6A modification is still poorly understood. For instance it is not clear what might be the role of the non-catalytic subunits of the KEOPS complex, since the mitochondrial Qri7 protein is capable on its own to catalyze the TCA transfer reaction onto tRNA. This may reflect requirements for t^6A biosynthesis that are different in the cytoplasm compared to those in the mitochondria. To answer these questions, structural data of KEOPS–tRNA complexes and more detailed reaction kinetics are needed. We determined the crystal structure of the Gon7/Pcc1 heterodimer and investigated the role of Gon7 within the KEOPS complex. We show that Gon7 interacts with Pcc1 and binds to Pcc1 at the opposite site of that between Pcc1 and Kae1, structurally mimicking the second Pcc1 subunit as in the archaeal Pcc1 homodimer. We further determined the structure of the yeast Cgi121/Bud32 subcomplex and construct a 3D model for the complete yeast KEOPS complex. This model revealed the presence of a deep conserved crater leading to the active site of Kae1 and surrounded by residues from Gon7, Pcc1, Kae1 and Bud32.

MATERIALS AND METHODS

Cloning, expression and purification of yeast KEOPS complex and subcomplexes

An updated sequence of the CGI121 gene published in 2007 contains a 106-bp intron (25). In order to remove this intron, the coding sequence has been cloned by using a 3 polymerase chain reaction (PCR) strategy with the following oligonucleotides (in bold is shown the NdeI restriction site on 'ex1_fw' and NotI restriction site on 'ex2_rv'; underlined is shown for the six histidine codons introduced at the 3′ end just before the Stop codon in order to graft a 6His-tag at the C-terminal end of Cgi121):

ex1_fw	5′-GGGAA**CATATG**GTAGTATCCATCATACCGC-3′
ex1_rv	5′-CTTAATTTATAGATTTTTCTTATGAGCGCTTC GTC-3′
ex2_fw	5′-AGCGCTCATAAGAAAAATCTATAAATTAAG TGATGA-3′
ex2_rv	5′-TTTTT**GCGGCCGC**TT AATGGTGATGGTGATGGTGCACACCCCTCAATT GAATAGC-3′

Using S288C *S. cerevisiae* genomic DNA as a template, the first exon of CGI121 gene was constructed by PCR using primers 'ex1_fw' and 'ex1_rv' while the second exon with primers 'ex2_fw' and 'ex2_rv'. Two microliter of each PCR products were then polymerized by overhang extension and used as a template in the third PCR carried out with primers 'ex1_fw' and 'ex2_rv'. Finally a 585-bp DNA fragment has been obtained, purified on agarose gel and double-digested with NdeI and NotI. This fragment has then been ligated into a pET9-derived expression vector previously digested with the same restriction enzymes. The ligation product has been used to transform XL1 blue MRF′ *E. coli* strain (Stratagen). After screening procedure, the selected pET9-Cgi121his plasmid has been sequenced on both strands to check for the absence of unwanted mutations.

We designed a 3.5-kb DNA synthetic fragment that contains five Open Reading Frames (ORFs) of *S. cerevisiae* KEOPS subunits and subcloned it into a pUC backbone vector (https://www.dna20.com/). In this expression vector named pJ241_KEOPS/EKC-WT, the five ORFs of KEOPS subunits have been placed in the following order from the 5′ end to the 3′ end: yKae1, yBud32, yCgi121, yPcc1 and yGon7–6His (26). Each ORF possesses its own RBS (ribosome binding site) and stop codon and only one T7 promoter and one terminator sequence have been added, respectively, at the 5′ and 3′ end of the synthetic gene. The

cloning of Bud32–Cgi121 complex was carried out using the same strategy as for KEOPS complex. Briefly, the new expression vector named 6-pK-Sc-Bu,Cg-a contains the upstream BUD32 and downstream CGI121 with six extra histidine codons at its 3′ end. In short, two other polycistronic vectors expressing Gon7-Pcc1-6His and Gon7-strep were chemically synthesized by genscript (Piscataway, NJ, USA) and subcloned into pET21a vector, which gave rise to the plasmids pET21a-BC1 and pET21a-BC4, respectively.

Cgi121-6His, KEOPS-6His and Bud32-Cgi121-6His were heterologously produced using *E. coli* BL21 Rosetta pLysS (Novagen), *E. coli* BL21 AI cells (Invitrogen) and *E. coli* BL21 Rosetta pLysS (Novagen) transformed with pET9-Cgi121his, pJ241_KEOPS/EKC-WT and 6-pK-Sc-Bu,Cg-a, respectively. The expression of KEOPS was induced by adding 0.2% of arabinose and isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM in 2YT media (Life Technologies), followed by growing the culture overnight at 15°C. The expression of Cgi121-6His and Bud32-Cgi121 was induced by adding IPTG (Sigma) to a final concentration of 0.5 mM for 3.5 h at 37°C in 2YT media (Life Technologies). Conventionally, SeMet-substituted Cgi121-6His was heterologously produced using *E. coli* BL21 Rosetta pLysS (Novagen) transformed with pET9-Cgi121-6His. The expression was induced at 15°C by adding IPTG to a final concentration of 0.5 mM for 72 h in a minimal media supplemented with amino acids and seleno-methionine. Native and SeMet-substituted Gon7-Pcc1-6His and N^{15}-labeled Gon7-strep were produced in *E. coli* RosettaBlue (DE3) pLysS Cells (Novagen) transformed with pET21a-BC1 and pET21a-BC4, respectively. The expression of native Gon7-Pcc1-6His was induced by adding IPTG to a final concentration of 0.5 mM in 2YT media (Life Technologies) overnight at 15°C. The expression of SeMet-substituted Gon7-Pcc1-6His and N^{15}-labeled Gon7-strep was induced by 0.5 mM IPTG in a minimal media supplemented with seleno-methionine and $^{15}NH_4Cl$, respectively, followed by growing the cells for 72 h at 15°C.

Cells were harvested by centrifugation and suspended in 30 ml of lysis buffer (20-mM Tris-HCl pH 7.5, 200 mM NaCl and 5 mM 2-mercaptoethanol) and either stored at −20°C or directly used for protein purification. Cells were first lysed by sonication (branson sonifier 250) and centrifuged at 20 000g for 30 min. For Cgi121-6His, KEOPS-6His, Bud32-Cgi121-6His and Gon7-Pcc1-6His, the supernatant (soluble fraction) was then applied on immobilized metal affinity chromatography (IMAC) using Ni-IDA resin (Macherey Nagel). Following a washing step with 25–30 ml of lysis buffer, elution of the five proteins was performed by loading successive fractions of 1 ml of lysis buffer supplemented with increasing concentrations of imidazole. A second size-exclusion chromatography step was performed on HiLoad Superdex 200 (for KEOPS-6His) or HiLoad Superdex 75 (GE-Healthcare) using preparative buffer comprising 20 mM Tris-HCl pH 7.5, 300 mM NaCl and 5 mM 2-mercaptoethanol. In some cases, an extra mono Q anion-exchange purification step was applied (buffers are composed of 30 mM Tris-HCl pH 8.5 and NaCl at a concentration of 100 mM and 1.0 M) in conjunction with a size-exclusion chromatography purification step. Appropriate fractions were pooled, concentrated, flash frozen and stored at −80°C or directly used for subsequent experiments. The purification of N^{15}-labeled Gon7-Strep used the Strep-Tactin resin (IBA), according to the manufacturer's instructions. The final size-exclusion chromatography step was performed on a HiLoad Superdex 75 column (20 mM potassium phosphate pH 4.5 and 200 mM NaCl). Fractions containing Gon7-Strep were pooled and prepared for immediate use.

Crystallization and structure determination

Both native and SeMet-substituted Cgi121 proteins were stored in 20 mM Tris-HCl pH 7.5, 300 mM NaCl, 5 mM 2-mercaptoethanol. Crystals were grown at 20°C from a 0.5:0.5 μl mixture of a 15-mg/ml protein solution with a crystallization solution composed of 1.2 M sodium citrate, 0.1 M sodium citrate pH 5**.** For data collection, the crystals were directly flash-frozen in liquid nitrogen. The native and SeMet diffraction data were recorded on beamlines ID14-EH1 and BM30A, respectively (ESRF, Grenoble, France).

The structure was determined by the SAD method (Single wavelength Anomalous Dispersion) using the anomalous signal from the selenium element. Data were processed using the XDS package (27). The space group was $P2_12_12_1$ with two copies per asymmetric unit. All the expected Se sites (five per monomer) were found with the program SHELXD using reflections in the 50–3 Å resolution range (28). Refinement of the Se atom positions, phasing and density modification were performed with the program SHARP (29). The quality of the experimental phases allowed automatic building of an almost complete model with the ARP/WARP application (30,31). This model was then refined against the 1.9 Å native data set using PHENIX (32) and then rebuilt with the 'TURBO' molecular modeling program (http://www.afmb.univ-mrs.fr/-TURBO). The final model contains residues 1–177 and 1–178 from monomers A and B, respectively. In addition, 383 water molecules could be modeled into the electron density maps. Statistics for all the data collections and refinement of the SeMet-substituted and native Cgi121 are summarized in Supplementary Table S1.

Bud32-Cgi121-6His was crystallized using the hanging drop vapor diffusion method. Crystals were obtained at 20°C from a crystallization solution that contains Bud32-Cgi121-6His at 5 mg/ml in 50 mM Tri-sodium pH 5.6, 10% 2-Propanol and 10% PEG4000 or 50 mM sodium acetate pH 4.6, 1 M ammonium sulfate. The crystals were cryo-protected by transfer into the crystallization solution with increasing glycerol concentrations up to 30% v/v and then flash cooled in liquid nitrogen. The diffraction data set was collected at SOLEIL Synchrotron beamline Proxima 1 (Saint Aubin, France). Data were processed using the XDS package (27). The space group was $I4_1$ with one heterodimer per asymmetric unit. The structure was determined by the molecular replacement method with the program MOLREP (33), using SeMet-substituted Cgi121 and *Methanocaldococcus jannaschii* Bud32 (15) structures as search models. The structure was further improved by the PHENIX AutoBuild wizard (34), refined against the 2.25 Å resolution native data set using REFMAC5 (35) and then

manually built with COOT (36). The final model for the Cgi121-Bud32-6His complex contains residues from 3 to 57 and 71 to 258 from Bud32 as well as residues 1 to 179 from Cgi121. We also observed well-defined electron densities for two glycerol moieties, two acetates and one adenosine molecule. In parallel, we cocrystalized Bud32-Cgi121-6His with various nucleotides and obtained crystals of complexes with $MgCl_2$ and AMPPNP, ATP and AMP. The crystals were grown under the same conditions as for the Bud32–Cgi121 apo-complex. These crystals belong to the $I4_1$ space group and diffracted up to 1.8 Å resolution at SOLEIL Synchrotron beamline Proxima 1 (Saint Aubin, France). Analysis of the Fourier difference map enabled us to clearly locate all the ADP-moieties of nucleotides that are bound in Bud32. However, the precise position of the γ-phosphate of ATP and AMPNP is unknown due to the lack of electron density. Statistics for all the data collections and refinement of the different structures are summarized in Table 1.

The freshly purified native or SeMet-substituted Pcc1-Gon7 complexes were concentrated to 25 mg/ml and crystallized using vapor diffusion. All the crystals were obtained at 18°C from crystallization reservoir that mixed an equal volume of the Pcc1–Gon7 complex with a solution containing 20% glycerol, 20% PEG1500, 0.1 M acetate sodium pH 4.2, 3% methanol and 0.5 M NaCl. Crystals were cryo-protected by soaking with crystallization liquor supplemented with 25% PEG400 v/v in liquid nitrogen prior to X-ray data collection. The native and SAD data sets were collected at 100 K on the SOLEIL Synchrotron beamline Proxima 1 (Saint Aubin, France). The native and SeMet-substituted data sets were processed at 3.0 Å and 2.5 Å, respectively, with the XDS package (37). Both complexes crystallized in the space group C2 with three copies of the Pcc1–Gon7 complex in the asymmetric unit (cell parameters are summarized in Table 1). The structure of the SeMet-substituted Pcc1–Gon7 complex was determined by the SAD method. Nine selenium sites (six from Pcc1 and three from Gon7) out of a total of 18 were located using the program SHELXD (38). The phasing and the refinement of the Se sites were performed with Phaser (39) and Parrot (40), respectively. The model building was carried out with the Buccaneer package (41) and the COOT (36) program, taking advantage of the structural model of *Pyrococcus furiosus* Pcc1 (17). The final model of the SeMet Pcc1–Gon7 complex was refined with REFMAC5 (35). The three copies of the complex in the asymmetric unit are identical and the model contains residues 9–90 of Pcc1 (native full length contains 1–88 plus 6His) and residues 3–21, 65–113 of Gon7 (native full length contains 1–123). The missing residues were not observed in any of the complexes in the asymmetric unit and are hence disordered. The structure of SeMet-substituted Pcc1–Gon7 complex was used to phase the native data set using MOLREP (42) and the final model of the Pcc1–Gon7 complex was built with COOT (36) and refined against the native data set using REFMAC5 (35). The data collection and refinement statistics are summarized in Table 1.

Size-exclusion-chromatography-coupled multi-angle laser light scattering

The determination of the molecular weight of the proteins in solution was carried out by size-exclusion-chromatography-coupled multi-angle laser light scattering (SEC-MALLS) analysis using a GPCMax-TDA system (Viscotek, Malvern, France). The Superdex TM 75 HR 10/30 or Superdex TM 200 HR 10/30 columns (GE Healthcare) for size-exclusion chromatography were equilibrated with buffer composed of 20 mM Tris pH 7.5, 200 mM NaCl and 5 mM 2-mercaptoethanol. Eighty to hundred microliters Gon7-Pcc1-6His or KEOPS-6His at a concentration of 3–5 mg/ml were injected at a flow rate of 0.5 ml.min^{-1} onto the SEC column and eluted with the equilibration buffer at a flow rate of 0.3 ml min^{-1}. Elution was followed by a UV-visible spectrophotometer, a differential refractometer, a 7° low angle light scattering detector, a 90° right angle light scattering detector and a differential pressure viscometer. The OmniSEC software program was used for the acquisition and analysis of the data. Bovine serum albumin was used as standard reference protein of known molecular weight, concentration, refractive index increment (dn/dc = 0.185 ml.g^{-1}), and intrinsic viscosity to determine the instrument response factors of the detectors for the mobile phase being used.

Small angle x-ray scattering

Small angle x-ray scattering (SAXS) experiments were carried out at the SOLEIL synchrotron SWING beamline (Saint-Aubin, France). The sample to detector (Aviex CCD) distance was set to 1500 mm, allowing reliable data collection over the momentum transfer range 0.008 Å^{-1} < q < 0.5 Å^{-1} with $q = 4\pi\sin\theta/\lambda$, where 2θ is the scattering angle and λ is the wavelength of the X-rays (λ = 1.0 Å). In order to separate the various species in solution, SAXS data were collected on samples eluting from an online size-exclusion high-performance liquid chromatography (SE-HPLCBio-SEC3 Agilent) column available on SWING and directly connected to the SAXS measuring cell. The purified KEOPS or Pcc1–Gon7 complexes were injected in the column pre-equilibrated with preparative buffer comprising 20 mM Tris pH 7.5, 200 mM NaCl and 5 mM 2-mercaptoethanol. Flow rate was 200 µl/min, frame duration was 1.0 s and the dead time between frames was 0.5 s. For each frame, the protein concentration (between 0.5 and 2.0 mg/ml at the top of elution peak) was estimated from UV absorption at 280 and 295 nm using a spectrometer located immediately upstream of the SAXS measuring cell. Selected identical frames corresponding to each elution peak were averaged. A large number of frames were collected before the void volume and averaged to account for buffer scattering. SAXS data were normalized to the intensity of the incident beam and background (i.e. the elution buffer) substracted using the programs FoxTrot (courtesy of SWING beamline) and Primus (43,44). The scattered intensities were displayed on an absolute scale using the scattering by water. In order to determine unambiguously the oligomeric state of the protein or complex, the molar mass was obtained using the macromolecule volume

Table 1. Data collection and refinement statistics of the Bud32/Cgi121 and Gon7/Pcc1 binary complexes

	Bud32–Cgi121 complex				Gon7–Pcc1 complex	
Data collection	Apo form	AMPPn[a]	ADP	AMP	Se-Met	Native
Wavelength (Å)	0.97918	0.98011	0.97911	0.97911	0.97895	0.98010
Resolution (Å)	43.20–2.95 (3.06–2.95)	43.71–2.0 (2.07–1.99)	43.62–1.95 (2.02–1.95)	44.01–1.67 (1.72–1.66)	43.53–2.99 (3.10–2.99)	44.28–2.44 (2.53–2.44)
Space group	I 41	I 41	I 41	I 41	C 121	C 121
Cell dimension						
a (Å)	111.54	112.93	113.01	113.89	79.10	80.73
b (Å)	111.54	112.93	113.01	113.89	102.50	104.39
c (Å)	86.42	87.23	86.35	87.41	88.80	89.89
α (°)	90.00	90.00	90.00	90.00	90.00	90.00
β (°)	90.00	90.00	90.00	90.00	111.80	111.55
γ (°)	90.00	90.00	90.00	90.00	90.00	90.00
Tot. reflections	51370	283335	236236	236981	51149	94329
Uni. reflections	11205	37120 (3660)	39488 (3892)	64291 (6337)	13327 (1325)	25517 (2428)
Completeness	99.8 (99.1)	99.79 (98.10)	99.89 (98.96)	99.41 (98.46)	99.54 (97.50)	98.9 (94.59)
Mean *I*/sigma (*I*)	14.32 (2.52)	27.29 (5.54)	23.72 (6.87)	18.48 (3.01)	8.08 (2.21)	14.69 (2.23)
R_{meas} (%)	12.4 (74.6)	4.9 (40.1)	5.4 (26.9)	4.5 (49.1)	11.2 (46.7)	7.6 (58.5)
CC (1/2) (%)	98.5 (82.0)	99.8 (98.8)	99.9 (97.4)	99.8 (86.4)	99.54 (97.50)	99.9 (87.8)
Refinement						
R_{factor} (%)	21.1	18.3	17.9	19.3	25.1	22.5
R_{free} (%)	28.7	22.2	20.9	21.3	30.8	26.9
Total atoms	2971	3460	3548	3423	3078	3204
Average B (Å^2)	25.70	34.0	31.0	29.0	66.0	64.0
RMSD from standard stereochemistry						
Bond length (Å)	0.010	0.010	0.005	0.008	0.011	0.014
Bond angles (°)	1.28	1.29	0.98	1.08	1.48	1.70
Ramachandran plot statistics						
Favored (%)	90	98	97	98	98.3	98.4
Allowed (%)	10	2	3	2	1.7	1.6
Disallowed (%)	0	0	0	0	0	0
PDB code	4WWA	4WW5	4WW9	4WW7	4WX8	4WXA

[a]The AMPPNP moiety without γ-phosphate.
The numbers in parentheses are for the highest resolution shell.

and the SAXS-MoW method available at http://www.if.sc.usp.br/~saxs/.

The model of *S. cerevisiae* Kae1 was obtained with Phyre^2 (45) using the crystal structures of *P. abyssi* Kae1 (12), *M. jannaschii* MJ1130 Kae1 (16), *S. typhimurium* YgjD (13) and *S. cerevisiae* Qri7 (8). The missing loop between residues 57 and 71 in crystal structure of Bud32 was also modeled and added by Phyre^2. The structural model of yeast KEOPS was obtained by replacing the archaeal KEOPS components with an *S. cerevisiae* Kae1 model combined with the crystal structures of the Bud32–Cgi121 and Pcc1–Gon7 complexes. The final model does not exhibit pronounced steric clashes at the subunit interfaces and is consistent with the mutations that disrupt the interactions between the subunits (17).

RESULTS

Crystal structure of the Gon7/Pcc1 complex

We made use of the ensemble of available structures of KEOPS subcomplexes to construct the complete archaeal complex showing that Pcc1, Kae1, Bud32 and Cgi121 are arranged in a linear manner. The Kae1 subunit is situated at the center of the complex, while Pcc1 binds to its N-terminal and Bud32 to its C-terminal lobe. Cgi121 binds to Bud32 at the opposite site of Kae1 (Supplementary Figure S1). We speculated that Gon7 would bind to one of the extremities of the yeast KEOPS complex, associating either with Pcc1 or with Cgi121. Deletion of the Gon7 gene is deleterious for cell life, but not the deletion of Cgi121, suggesting that Gon7 might interact with Pcc1 rather than with Cgi121. We therefore first co-expressed Gon7 and His-tagged Pcc1 from a polycistronic plasmid and observed that both proteins co-purified and form a stable complex (Supplementary Figure S2C). We obtained good quality preparations of the Gon7/Pcc1 complex that yielded diffracting crystals. The X-ray crystal structure of the Pcc1/Gon7 complex was determined using Se-methionine labeling. There are three copies of the Gon7/Pcc1 heterodimer in the asymmetric unit. Gon7 is partially ordered since electron density could only be observed for the peptide regions comprised between residues 3 to 21, 63 to 113. The missing regions are about the same for the three copies of Gon7 in the asymmetric unit, suggesting that the lack of electron density is due to disorder. Gon7 is made of an anti-parallel two-stranded β-sheet (residues 3–21) and a long α-helix (residues 63–95) that are packed to form the structured core of the protein (Figure 1A). Two copies of Gon7 in the asymmetric unit have a supplementary α-helix (residues 98–113) that is pointing away from the structured core, contacting the helical region of a neighboring Gon7/Pcc1 heterodimer in the crystal (not shown). The structure of Pcc1 (88 residues) is well defined

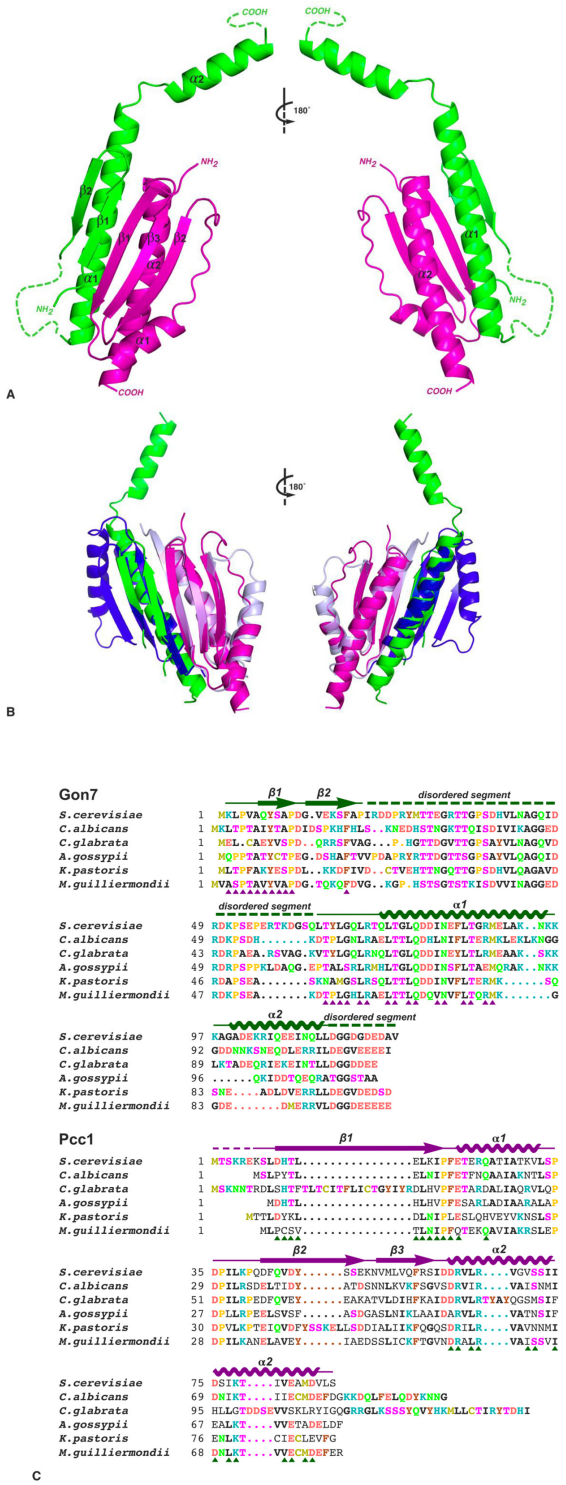

Figure 1. Crystal structure of the Gon7/Pcc1 complex from yeast. (**A**) The cartoon representation of the complex: Gon7 colored in green and Pcc1 colored in magenta. The secondary structure elements are labeled as for panel (**C**). Tthe principal missing region in the structure of Gon7 is shown as a green dashed line. (**B**) The structural superposition of the *P. furiosus* Pcc1 homodimer (the two subunits are colored in blue and gray, PDB code: 3ENC) onto the Gon7/Pcc1 complex (same color code as in panel (A)). (C) Sequence alignments of a few fungal Gon7 and Pcc1 sequences. The secondary structure elements as extracted from the crystal structure are superposed. Residues that are responsible for heterodimer contacts are indicated by triangles.

between residues 8 (or 9, depending on the copy) and 91 (three histidines from the 6His tag have defined electron density). Pcc1 forms a three-stranded anti-parallel β-sheet with two helices packed on one face. The structure of yeast Pcc1 is very similar to that of *P. furiosus* Pcc1 (17), with a Z-score of 7.7 and an RMSD of 1.09 Å for 65 aligned C_α atoms. Dimer formation between Gon7 and Pcc1 involves packing of two helices and formation of an anti-parallel β-sheet. The first β-strand of Gon7 aligns with the first β-strand of Pcc1, creating a continuous anti-parallel five-stranded β-sheet (Figure 1A, left panel). Complex formation buries about 2400 $Å^2$ for a total of 11 700 $Å^2$ accessible surface area. The most important contribution to the dimer interface results from the packing of the α1-helix of Gon7 against the α2-helix of Pcc1 (Figure 1A, right panel). Fifteen hydrogen bonds and five salt bridges stabilize the interface. The well-conserved residues from the Gon7 α1 helix directly interact with Pcc1: hydrophobic residues form the core of the interface while polar and charged residues form specific hydrogen bonds at the periphery (for instance the totally conserved Arg^{72} of Gon7 H-bonds to Glu^{83} of Pcc1). As shown in Figure 1C, the sequence of the β1 strand of Gon7 is not well conserved but the majority of its interactions with Pcc1 consist of hydrogen bonds between main chain atoms. The main non-structured region of Gon7 is situated between β2 and α1 and is not close to the heterodimer interface. Despite its disorder, this region contains a few very well-conserved sequence stretches and is globally better conserved than the interacting β1 strand. This suggests that this region might be functionally important and might be involved in interaction with other KEOPS partners or substrate tRNA. The very acidic C-terminal peptide of Gon7 is also disordered.

Archaeal Pcc1 alone or in complex with Kae1 forms a tight homodimer, characterized by a continuous six-stranded anti-parallel β-sheet and packing of two helices. Pcc1 dimer formation was also observed in solution and it was suggested that Pcc1 functions as a dimerization module of the KEOPS complex in archaea (Supplementary Figure S1). Modeling shows that it is, in principle, possible to construct a Pcc1 homodimer (not shown). We were however not able to express Pcc1 alone in soluble form and could therefore not test whether Gon7 competes with Pcc1 homodimer formation *in vitro*. Superposition of the Gon7/Pcc1 complex onto the Pcc1 homodimer reveals that their interfaces are very similar (Figure 1B). The structured part of Gon7 overlaps very well with the secondary structure elements of Pcc1 that are involved in homodimer formation. It should be noted that, although Gon7 and Pcc1 share the same architecture, they do not have the same fold. We conclude that Gon7 is a structural mimic of Pcc1 and precludes formation of the Pcc1 homodimer.

Gon7/Pcc1 in solution

Sequence-based structure prediction indicates that Gon7 is almost fully disordered (Supplementary Figure S2A). Since more than half of the Gon7 protein was not observed in the crystal, we investigated whether Gon7 becomes partially structured upon complex formation with Pcc1. We therefore prepared N^{15}-labeled Gon7 and investigated the folding of

the construct by nuclear magnetic resonance (NMR). The HSQC spectrum of Gon7 shows that the protein alone in solution is unstructured (Supplementary Figure S2B). We attempted to titrate N^{15}-labeled Gon7 with unlabeled Pcc1, but the latter protein forms stable soluble high molecular weight aggregates when produced alone and hence did not affect the HSQC spectrum of Gon7 (data not shown).

We then examined the oligomerization state of the Gon7/Pcc1 complex in solution by SEC-MALLS. The extracted molecular weight of Gon7/Pcc1 (34.6 kDa) was significantly larger than expected for a heterodimer (24.4 kDa; Supplementary Figure S2D). We noticed the presence of two types of hetero-tetramer in the crystal packing of Gon7/Pcc1, each composed of two Gon7 and two Pcc1 subunits. In hetero-tetramer type 1, two central Gon7 proteins align their beta sheets forming a continuous 10-stranded beta sheet with the Pcc1 units (Supplementary Figure S3B). The packing of the helical surface of two heterodimers created the heterotetramer type 2 (Supplementary Figure S3D). The main contribution for the interaction between the heterodimers comes from the C-terminal helix of Gon7 that swaps over to the long helix of Pcc1 in the other heterodimer. The SEC-MALLS molecular weight is in between those expected for a heterodimer and for a heterotetramer, suggesting they may be in dynamic equilibrium.

We examined the structure of the Gon7/Pcc1 complex in solution by SAXS. The X-ray scattering curves are presented in Supplementary Figure S3A and structural parameters summarized in Supplementary Table S2. The experiments were performed using a gel filtration column coupled to the SAXS instrument (46). SAXS data were collected online throughout the elution time. Selected identical scattering curves corresponding to the start of the elution peak were averaged. The value of the molar mass M obtained from the complete $I(q)$ curve ($M = 51 \pm 5$ kDa) is compatible with the calculated mass of the Gon7/Pcc1 heterotetramer ($M_{calc} = 48.8$ kDa). Unfortunately, the SAXS data do not allow choosing between the two types of heterotetramers (Supplementary Figure S3B and D), as both can be equally well incorporated within the envelope of the complex deduced from the SAXS curve using the program GASBOR (47).

Crystal structure of the binary complex of Bud32/Cgi12 bound to ADP

We expressed and purified the yeast Bud32/Cgi121 complex, and determined its crystal structure in complex with ADP (Figure 2). Crystal structures of archaeal and human Cgi121 and the archaeal Kae1/Bud32/Cgi121 ternary complex were previously reported (17). The structure shows that yeast Cgi121 interacts with the N-terminal lobe of Bud32 and that its structure is very similar to that of its human and archaeal homologs (RMSD of 0.6 Å for all superposed C_α atoms). Yeast Cgi121 is made of a central four-stranded β-sheet and eight α-helices. Six helices group into a globular bundle on one side of the β-sheet and provide the main interaction site with Bud32. Archaeal Cgi121 has a shorter N-terminal region compared to the human and yeast orthologs and therefore its central β-sheet only contains three strands compared to four in the eukaryotic structures. The structural comparison between free and Bud32-bound Cgi121 shows that they are very similar except for the segment between the α1-helix and β3-strand. This segment is structured as a loop in the Bud32-bound form Cgi121 while it forms a one helical turn in the apoform Cgi121 (Supplementary Figure S4A). The N-terminal lobe of Bud32 interacts with Cgi121 and the C-terminal lobe interacts with Kae1 as shown in crystal structures of Kae1/Bud32 and Kae1/Bud32/Cgi121 archaeal complexes (16,17). The fourth β-strand from the N-terminal lobe of the Bud32 fits into the space created by the parallel α2- and α6-helices from Cgi121. The structure of the binary yeast Bud32/Cgi121 complex is similar to that described in the context of the archaeal Kae1/Bud32/Cgi121 ternary complex (17) but with a few pronounced differences (Supplementary Figure S4). First, additional contacts are made between Bud32 and Cgi121 in the yeast complex involving residues 4–10 from Cgi121 (connecting strands 1 and 2) and residues 35–42 from Bud32. The latter connection has an insertion of 11 residues compared to archaeal homologs. The contact surface of the yeast Bud32–Cgi121 complex (about 1175 $Å^2$) is therefore larger than for its archaeal ortholog (700 $Å^2$). The Cgi121–Bud32 interface is composed of both hydrophobic and polar interactions (nine hydrogen bonds). A second important structural difference with the archaeal complexes resides in Bud32. Bud32 is a small atypical protein kinase with a shortened C-lobe lacking the activation segment. The structure of yeast Bud32 is well defined in the complex, except for the disordered connection (residues 58–68) between the third β-strand and the second α-helix. We were unable to obtain crystals for yeast Bud32 alone and hence ignore whether conformational changes are induced by Cgi121 binding. We compared the crystal structures of Bud32 as present in yeast Bud32/Cgi121 with those from the archaeal binary and ternary complexes. Generally, the N-terminal lobe of protein kinases is composed of five strands and a flanking α-helix. The first strand of Bud32 was found to be disordered in all reported structures of archaeal Bud32 but is well defined in the yeast Bud32/Cgi121 complex. The N-terminus of Bud32 forms a short α-helix followed by a loop that contacts Cgi121 and that is absent in its archaeal homologs (Supplementary Figure S4B). Although both the N- and C-terminal lobes of yeast and archaeal Bud32 are superposed well individually, their relative orientations are different. Upon superposition of the N-terminal lobes of yeast and archaeal Bud32, their C-terminal lobes are slightly shifted (data not shown).

The structure of the *M. jannaschii* Kae1/Bud32 complex was determined in the presence of the nucleotide analog AMPPNP. Although the electron density for AMPPNP was clearly defined in the active site of Kae1, it was absent from Bud32 (16). Likewise, no binding of nucleotide was observed in Bud32 in the structure of the ternary Kae1/Bud32/Cgi121 complex from archaea (17). We cocrystallized yeast Bud32/Cgi121 in the presence of ATP, ADP, AMPPNP and AMP and in all cases the nucleotide was well bound at the active site of Bud32 (Supplementary Figure S5). The adenosine base occupies a hydrophobic pocket created by residues of the linker region between the two lobes. For all nucleotides the α- and β-

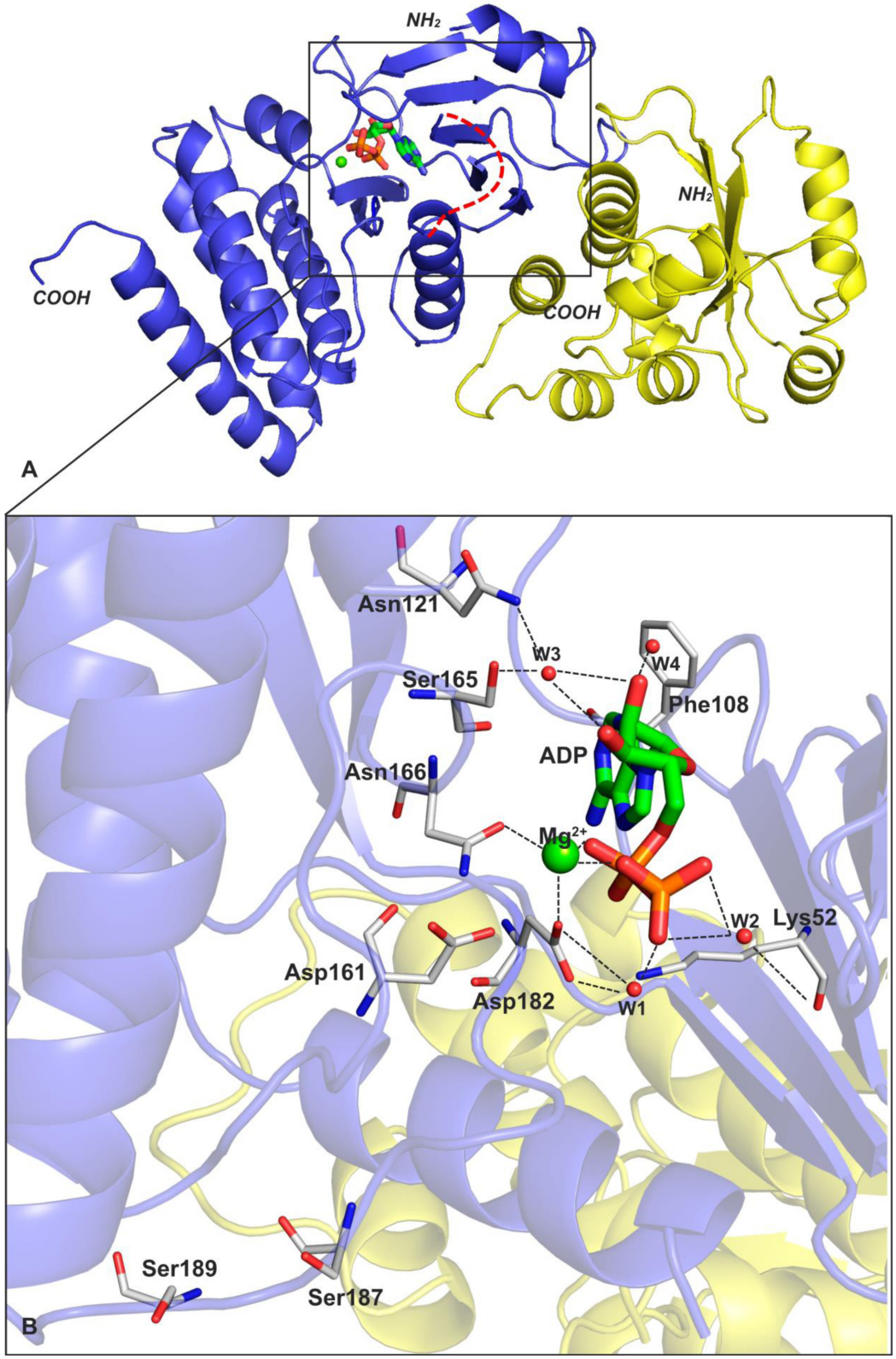

Figure 2. The structure of Bud32/Cgi121 binary complex from yeast. (**A**) Cartoon representation showing Cgi121 (colored in yellow) binding to the N-lobe of Bud32 (colored in blue). ADP (thick sticks) and Mg^{2+} (green sphere) are bound in the active site between the two lobes of Bud32. The red line denotes the missing loop between Arg^{58} and Leu^{68}. (**B**) The binding of ADP and Mg^{2+} in the active site of Bud32. Residues whose mutation abolished the kinase (K52, D161 and D182), autophosphorylation (S187 and S189) and ATPase activities (D161) of Bud32 are represented in sticks and labeled. The water molecules are represented as red spheres and labeled as W with numbering.

phosphates are well defined and in some cases bound to a divalent Mg^{2+} ion, which Asp^{182} does tightly coordinate. The α- and β-phosphate groups further interact with Lys^{52}, a very conserved residue in the active site of kinases (Figure 2B and Supplementary Figure S6). Lys^{52} is involved in the stabilization of the penta-covalent transition state during the phosphorylation reaction. The metal ion is further liganded to the side chains of Asn^{166} and the conserved Asp^{182} (Figure 2 and Supplementary Figure S6). The P-loop between the β1 and β2 strands of Bud32 (Ser^{23}-Ile^{29}) was disordered in structures of archaeal Bud32 in the absence of nucleotides but has a well-defined conformation in yeast Bud32 in the presence of nucleotides. We did not observe electron density for the γ-phosphate for any of the nucleotide complexes. We presume the γ-phosphate is disordered since it was also not observed in the non-hydrolysable AMPPNP analog. The presence of the nucleotide in the active site of the yeast Bud32 structure is coherent with the ordered β1-strand and P-loop. The catalytic cleft residues whose mutations destroy the auto-phosphorylation activity of Bud32 and cause cell growth and telomere length defects (Lys^{52}, Asp^{182}, Asp^{161}) are all directly bound to ADP in Bud32 (16,17,48). Also, Ser^{187} and Ser^{189} that are autophosphorylation sites of Bud32 reside in the loop, which is accessible to the ADP binding site (Figure 2) (48).

Model of the complete yeast KEOPS complex

A model of archaeal KEOPS was previously proposed based on the structures of subcomplexes (17). We used this approach to construct a model of yeast KEOPS by superposing the present Bud32/Cgi121 and Pcc1/Gon7 structures onto their orthologs in archeal KEOPS. We incorporated a yeast Kae1 model that was obtained by the Phyre^2 web server (45). The model for the complete yeast KEOPS complex is represented in Figure 3A. Like archaeal KEOPS, yeast KEOPS forms a linear arrangement of subunits, with Gon7 at one end and Cgi121 at the other. The active site of Kae1, centered on the conserved metal cluster and containing the ATP binding site, is at the bottom of a crater whose walls are made of Kae1, Bud32, Pcc1 and Gon7.

The crystal structures of archaeal Pcc1 alone and of the Pcc1/Kae1 complex showed that Pcc1 forms homodimers (17). It was therefore suggested that the intact archaeal complex would form a dimer of heterotetramers for which Pcc1 would function as the dimerization unit (Supplementary Figure S1) (9). The structure of the Gon7/Pcc1 dimer shows that Pcc1 occupies the same interface as that used for Pcc1 homodimer formation (Figure 1B and C). We therefore set out to determine the quaternary structure of the complete yeast KEOPS complex. Recombinant yeast KEOPS complex was purified after high yield expression from a polycistronic construct in *E. coli* cells. We first determined the molecular mass in solution of the recombinant yeast KEOPS complex by SEC-MALLS. The extracted molecular weight was 109.7 ± 5 kDa (Figure 4A), which is close to the theoretical molecular weight of the pentameric KEOPS complex (117.7 kDa), suggesting yeast KEOPS forms a hetero-pentamer in solution. In order to confirm the stoichiometry and to evaluate our model of the complex, we investigated the intact KEOPS complex in solution by SAXS coupled to gel filtration. The molar mass extracted from the scattering curve corresponding to the top of the elution peak is equal to 122 ± 10 kD (Supplementary Table S2), in good agreement with the SEC-MALLS results. We evaluated our above-mentioned linear model of the yeast KEOPS complex against the SAXS data. The program BUNCH (49) was used for modeling the missing parts of the model (mainly in the Gon7 subunit) as chains of dummy residues whose C-alpha positions are separated by 3.8 Å. Reconstructing the backbone and adding the side chains using the program PD2 (50) then completed the model. The scattering amplitude curves calculated on the final model by using the program CRYSOL (51) is in good agreement with the experimental curve ($\chi = 1.25$) (Figure 4B). These data validate the quaternary structure and the molecular architecture of our model of yeast KEOPS.

It was shown that the five KEOPS subunits are required for the biosynthesis of tRNA t^6A (6,8) and that Kae1 transfers the threonylcarbamoyl moiety from TCA onto substrate tRNA (6,9). In our model, the catalytic center of Kae1 contacts the C-terminal region of Bud32 and Pcc1 and Gon7 surround the active site entrance. The bottom of the crater formed by Gon7-Pcc1-Kae1-Bud32 has a positively charged and conserved surface (Figures 3 and 5). After refinement of the Bud32-Cgi121 structure we observed residual electron density for four sulphate ions. In a previously determined structure of *P. abyssi* Kae1 we identified three tungstate ions (12) after soaking the crystals in a tungstate solution for phasing the diffraction data. Sulfate and tungstate ions often occupy the positions of DNA- or RNA-backbone phosphates in apo-crystal structures of nucleic acid binding proteins. We positioned these sulfate and tungstate ions onto our model of the yeast KEOPS complex in Figure 5. One sulfate ion occupies an ATP phosphate position in the active site of Bud32 and two others are bound at the end of its C-terminal helix, close to the entrance of the active site of Kae1. When the structure of *P. abyssi* Kae1 is superposed onto the modeled yeast ternary complex, two of the tungstate ions are found close to the C-terminal of Bud32. The presence of these ions and the surrounding positive surface patch suggest that this region of the complex could be involved in the binding of nucleic acids. We generated a model for the KEOPS–tRNA (PDB: 3L0U) complex by rigid global docking with PatchDock (53), specifying that A37 of tRNA directly contacts His^{141}/His^{145}/Asp^{344}. The resulting model orients the anticodon stem loop of tRNA toward the active site of Kae1 and shows that all the subunits except Cgi121 of KEOPS are potentially involved in the binding of tRNA (Figure 5). We realize that tRNA probably undergoes important structural rearrangements during the modification reaction, precluding more rigorous modeling.

DISCUSSION

Experimental data obtained from archaeal and eukaryotic t^6A biosynthesis systems led to the conclusion that the KEOPS complex catalyzes the transfer of the threonylcarbamoyl moiety from the unstable TCA intermediate onto the A_{37} base of acceptor tRNA (7–9). The active sites of the t^6A biosynthesis systems are located on the totally con-

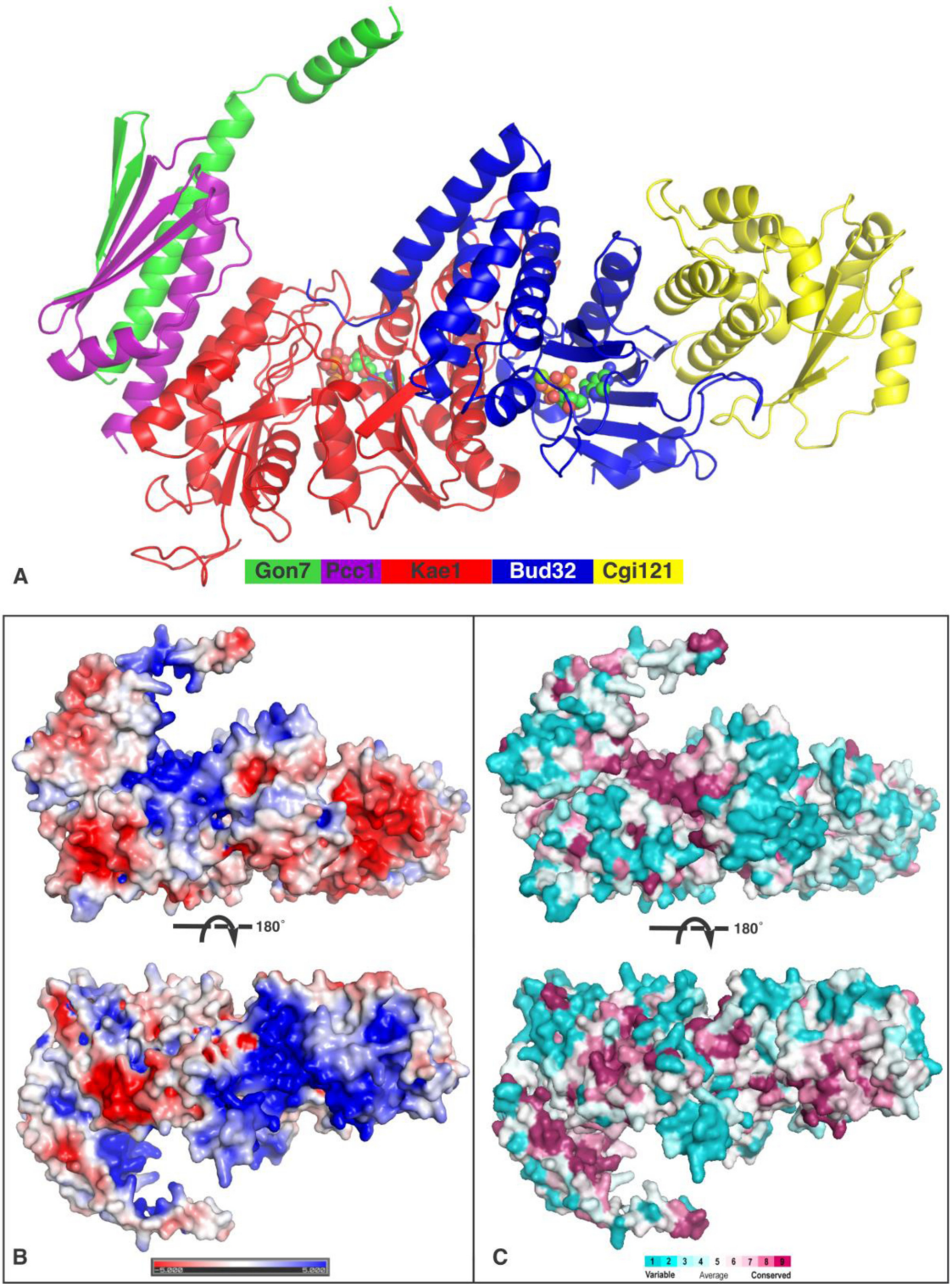

Figure 3. The reconstituted model of yeast KEOPS. (**A**) Cartoon representation of yeast KEOPS: Gon7 (green), Pcc1 (purple), Kae1 (red), Bud32 (blue) and Cgi121 (yellow). The ATP and ADP moieties in Kae1 (modeled from *P. abyssi* Kae1, PDB: 2IVP) and Bud32 are shown as spheres. (**B**) The electrostatic potential surface. The potential is given with the negative (red) and positive (blue) contour levels in the range from –5.0 to +5.0 kBT, respectively. (**C**) The sequence conservation projected onto the surface as reported by ConSurf (http://consurf.tau.ac.il/). The least conserved residues and highest conserved residues are scaled from cyan to dark purple. All three structure representations are rendered in the same orientation.

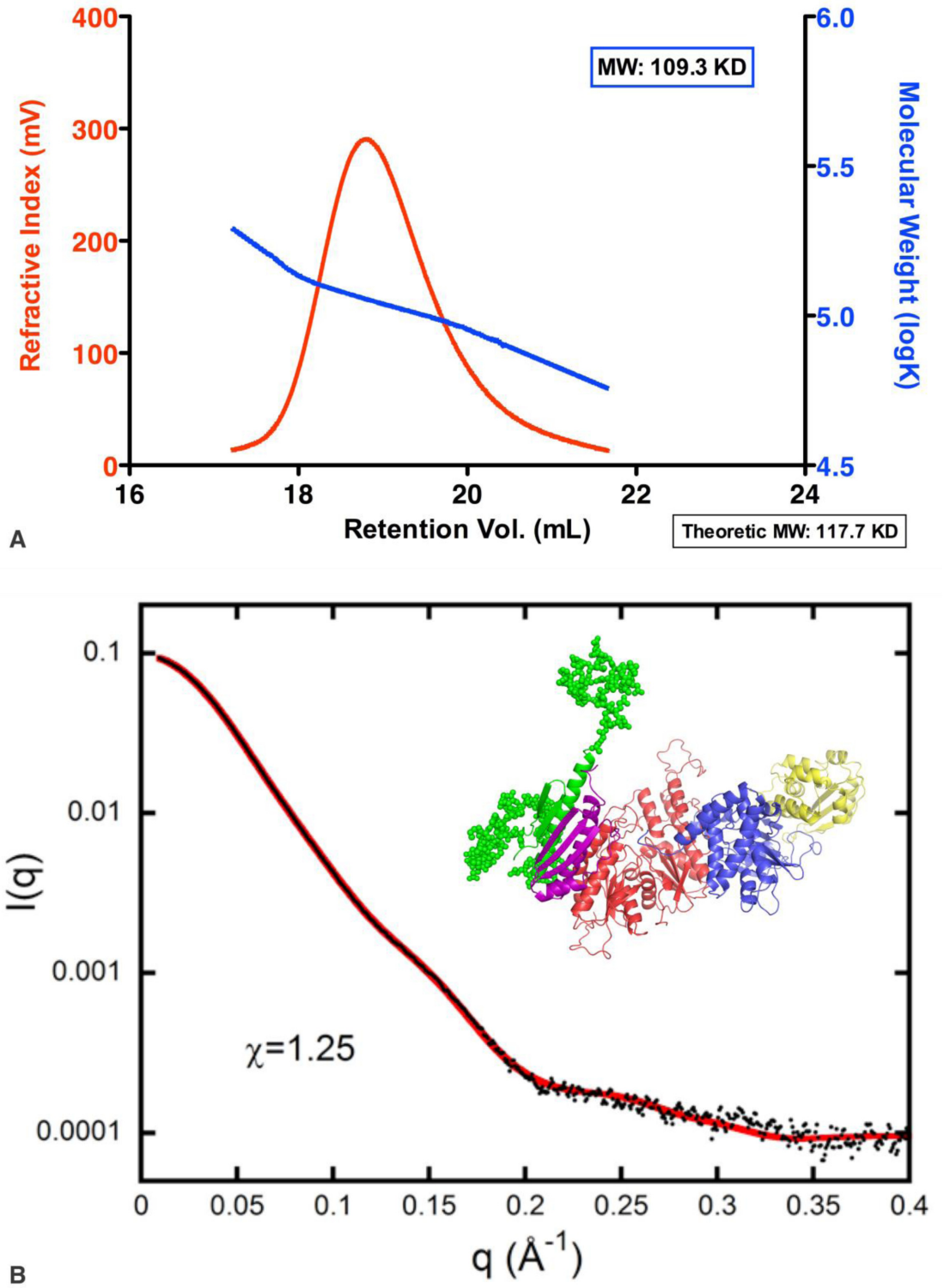

Figure 4. The determination of the quaternary structure of yeast KEOPS in solution. (**A**) The gel filtration profile and molecular weight extracted by SEC-MALLS. (**B**) The fitting of the yeast KEOPS model (red line) to the SAXS experimental curves (black dots). The missing regions of Gon7 were added by modeling with BUNCH program and represented as green spheres.

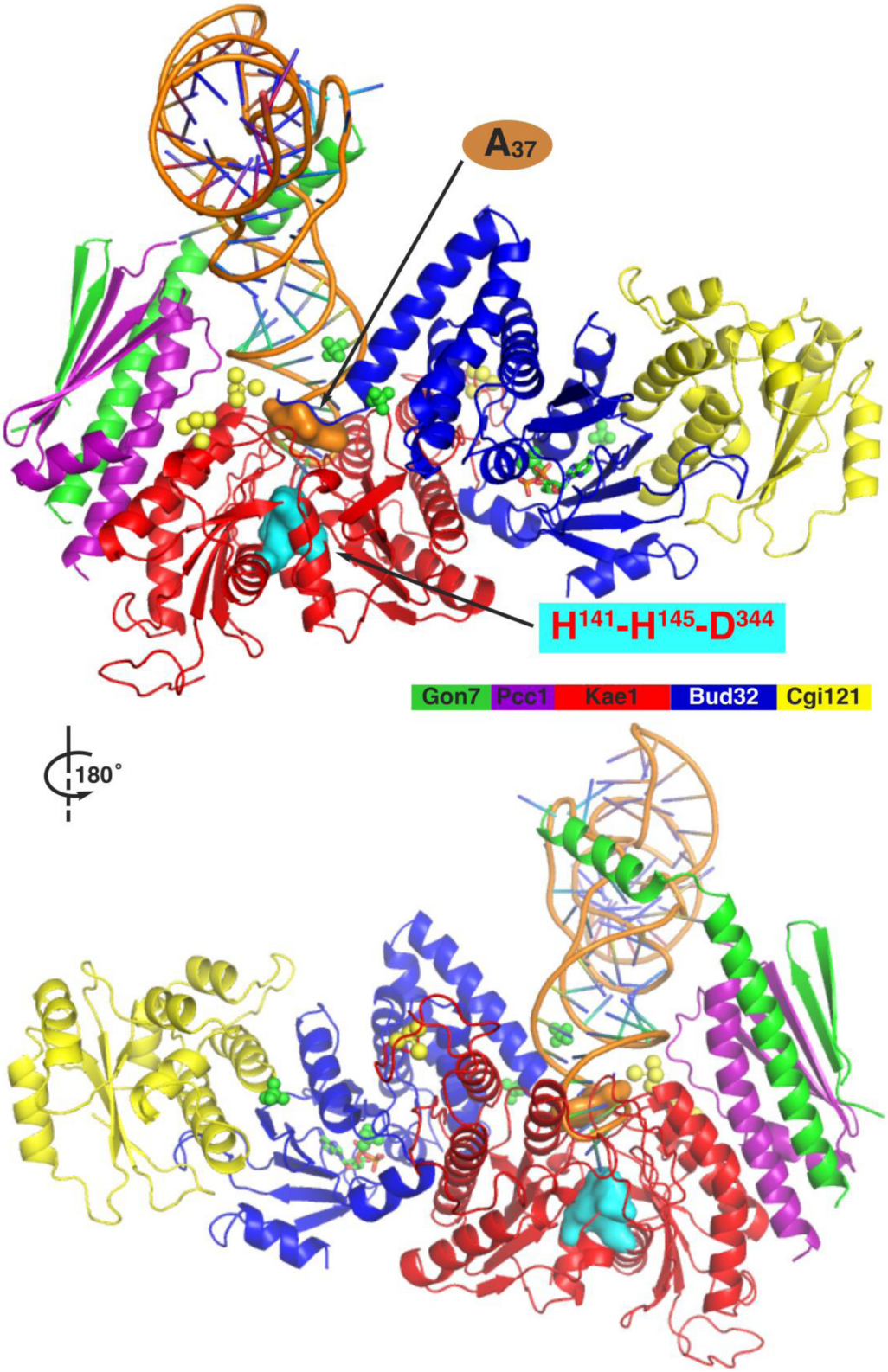

Figure 5. Docking model for KEOPS–tRNA complex by PatchDock. The A_{37} of tRNA and active site (H141, H145 and D344) of Kae1 are highlighted and denoted. The orientation is the same as for Figure 3. We have also represented the sulfate ions that were present in the yeast Bud32-Cgi121 structure (green spheres) and the tungstate ions present in the *P. abyssi* Kae1 structure (yellow spheres taken from PDB: 2IVO) were superposed onto the yeast Kae1 model.

served metal cluster in the Kae1/Qri7/YgjD proteins. However, the role of the other subunits (Bud32, Pcc1, Cgi121, Gon7) remains largely unknown. Analysis of the archaeal KEOPS complex showed that Pcc1, Bud32 and Kae1 are the minimal protein set for the *in vitro* t^6A biosynthesis but that Cgi121 increases the efficiency of the reaction (9). Bud32 was shown to possess both ATPase and kinase activities (9,17). Active site mutations of Bud32 (D161A/R in yeast and D127R in archaea) affect kinase activity (16,48), t^6A biosynthesis activity (6,9) and ATPase activity (9). The transfer of the threonylcarbamoyl moiety from TCA onto A_{37} of tRNA does not, in principle, require ATP hydrolysis nor the kinase activity of Bud32. This was clearly demonstrated by the study of the mitochondrial Qri7 enzyme, which is active without assistance of other proteins, forming together with Sua5 the minimal t^6A biosynthesis system. Though AMPPNP was present in the crystallization conditions, no nucleotide was bound in Bud32 in the structures of archaeal Bud32 (16,17), which was explained by its disordered N-terminus and P-loop region. ATP is firmly bound in the yeast Bud32/Cgi121 complex interacting in a canonical manner with the well-structured P-loop. Kae1 and Cgi121 seem to have antagonistic effects on the ATPase activity of Bud32. Bud32 in complex with Cgi121 was also shown to have autophosphorylation activity (54), which was inhibited by increasing concentrations of Kae1 (16).

Bud32 belongs to the RIO-type family of atypical small kinases that lack the kinase activation loop (55). Rio1 and Rio2 are involved in the maturation of the ribosome. It was recently shown that Rio2 does not only function as a kinase but that it exhibits robust ATPase activity *in vitro* (56). Interestingly, the energy released by ATP hydrolysis is required for dissociation of Rio2 from the ribosome, a necessary step in pre-40S particle maturation. The crystal structure of *Chaetomium thermophilum* Rio2 contained a phospho-aspartate intermediate typically found in P-type ATPases but unusual for protein kinases. Phosphoryl-transfer from ATP to Asp257 in Rio2's active site and subsequent hydrolysis of the aspartylphosphate was proposed to be a trigger to power 40S subunit biogenesis (56). Asp257 in *C. thermophilum* Rio2 is conserved in the Bud32 active site (corresponding residues are Asp142 in *P. abyssi* and Asp182 in *S. cerevisiae*) suggesting that also Bud32 could function as a P-type ATPase during the t^6A synthesis reaction. We hypothesize that Bud32 might regulate the interaction of the KEOPS complex with substrate tRNA. The positioning of the Bud32 subunit within our model of the KEOPS complex supports this hypothesis. The presence of the C-terminal helix of Bud32 near the active site of Kae1 offers an opportunity to couple ATP hydrolysis by Bud32 and tRNA binding by the KEOPS complex. The conserved and very positively charged C-terminus of Bud32 lies at the entrance of the active site of Kae1 (Figure 3B and Supplementary Figure S4). We therefore expect that Bud32 might interfere with the binding of the bulky tRNA substrate in the Kae1 active site. ATPases often undergo conformational transitions between the ATP and ADP bound states. ATP hydrolysis could influence the relative position of the N-terminal and C-terminal lobes of Bud32 and hence change, for instance, the position of the C-terminal helix, which in turn may modify its interactions with the substrate tRNA during t^6A biosynthesis. Many RNA-binding processes are regulated by nucleotide triphophosphate hydrolysis. For instance, the elongation factor EF-Tu undergoes a conformational change upon GTP hydrolysis that dissociates its complex with aminacyl-tRNA during peptide bond formation on the ribosome (57).

Mao *et al.* used the structure of archaeal Kae1/Bud32/Cgi121 complex to study the functional effects of complex disrupting mutations in yeast (17). Interface mutations (L172E/A176R) that interrupt interaction between Cgi121 and Bud32 displayed slow growth phenotype, analogous to the CGI121 gene deletion mutant. The present structure of the yeast Bud32/Cgi121 complex confirms that the interfacial residues targeted in Mao's study make crucial interactions in the yeast complex and hence these functional data are in full agreement with our structure and model.

Gon7, the fifth fungus-specific KEOPS subunit, was predicted to be largely unstructured (Supplementary Figure S2A), which we confirmed by NMR measurements (Supplementary Figure S2B). We showed here that it forms a tight complex with Pcc1, which functions as a dimerization unit of the archaeal KEOPS complex. Also the human KEOPS complex was reported to form a dimer of heterotetramers (58). Upon complex formation with Pcc1, about half of Gon7 acquires a well-defined structure and forms a continuous β-sheet structure with Pcc1. The binding of Gon7 to Pcc1 is incompatible with Pcc1 homodimer formation. Pcc1 can therefore not function as a dimerization unit in the yeast KEOPS complex, which we experimentally confirmed by Sec-MALLS and SAXS data (Figure 4).

Our model of the complete yeast KEOPS, which is compatible with our SAXS data (Figures 3 and 4), has all subunits aligned with Gon7 on one end and Cgi121 at the other. Our model shows that the region of the KEOPS complex around the ATP binding site of Kae1 is very positively charged (Figure 3B) and compatible with tRNA binding. Residues from Bud32, Pcc1 and Kae1 contribute to this positive surface patch, and the three subunits might be involved in tRNA binding (Figure 5). It should be mentioned that Qri7 forms homodimers and that the homodimer mimics the archaeal Pcc1/Kae1 and bacterial YgjD/YeaZ heterodimers. Disruption of the Qri7 homodimer results in loss of t^6A activity. The functional necessity of this higher order structure for t^6A activity is not yet clear, but we hypothesize that it is important for tRNA binding. The role of the fungi-specific Gon7 remains intriguing. We showed here that it imposes a different quaternary structure of the KEOPS complex adding another degree of complexity to the t^6A mechanism.

CONCLUSION

Based on the present crystal structures of Pcc1/Gon7 and Bud32/Cgi121. We propose a linear heteropentameric structure of the complete yeast KEOPS complex that is compatible with SAXS data. Our KEOPS model displays a conserved and positively charged surface formed by Bud32, Kae1 and Pcc1, suggesting they form the binding site for tRNA during t^6A biosynthesis. Further experimental investigation on the assembly of the KEOPS and tRNA binding

is needed to unravel its mechanism of tRNA t^6A biosynthesis and clarify its relationship with telomere elongation and transcriptional regulation.

ACCESSION NUMBERS

4WWA, 4WW5, 4WW9, 4WW7, 4WX8 and 4WXA.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Charlotte Saint-André for technical assistance in protein preparation; Andy Thompson (SOLEIL, St Aubin) for his assistance with the SAD data collection and Ewen Lescop (ICSN, Gif s/Yvette) for the NMR contribution.

FUNDING

The French Infrastructure for Integrated Structural Biology (FRISBI) [ANR-10-INSB-05-01]; the 'Lidex-BIG' Project Funding [IDEX Paris-Saclay ANR-11-IDEX-0003-02]; 2009 CSC Ph.D. Fellowship [to W.Z.].
Conflict of interest statement. None declared.

REFERENCES

1. Takemura,S., Murakami,M. and Miyazaki,M. (1969) The primary structure of isoleucine transfer ribonucleic acid from Torulopsis utilis. Complete digestion with ribonucleases and construction of model of its secondary structure. *J. Biochem.*, **65**, 553–566.
2. Parthasarathy,R., Ohrt,J.M. and Chheda,G.B. (1974) Conformation and possible role of hypermodified nucleosides adjacent to 3′-end of anticodon in tRNA: N-(purin-6-ylcarbamoyl)-L-threonine riboside. *Biochem. Biophys. Res. Commun.*, **60**, 211–218.
3. Murphy,F.V.T., Ramakrishnan,V., Malkiewicz,A. and Agris,P.F. (2004) The role of modifications in codon discrimination by tRNA(Lys)UUU. *Nat. Struct. Mol. Biol.*, **11**, 1186–1191.
4. El Yacoubi,B., Hatin,I., Deutsch,C., Kahveci,T., Rousset,J.P., Iwata-Reuyl,D., Murzin,A.G. and de Crecy-Lagard,V. (2011) A role for the universal Kae1/Qri7/YgjD (COG0533) family in tRNA modification. *EMBO J.*, **30**, 882–893.
5. Daugeron,M.C., Lenstra,T.L., Frizzarin,M., El Yacoubi,B., Liu,X., Baudin-Baillieu,A., Lijnzaad,P., Decourty,L., Saveanu,C., Jacquier,A. *et al.* (2011) Gcn4 misregulation reveals a direct role for the evolutionary conserved EKC/KEOPS in the t6A modification of tRNAs. *Nucleic Acids Res.*, **39**, 6148–6160.
6. Srinivasan,M., Mehta,P., Yu,Y., Prugar,E., Koonin,E.V., Karzai,A.W. and Sternglanz,R. (2011) The highly conserved KEOPS/EKC complex is essential for a universal tRNA modification, t6A. *EMBO J.*, **30**, 873–881.
7. Perrochia,L., Crozat,E., Hecker,A., Zhang,W., Bareille,J., Collinet,B., van Tilbeurgh,H., Forterre,P. and Basta,T. (2013) In vitro biosynthesis of a universal t6A tRNA modification in Archaea and Eukarya. *Nucleic Acids Res.*, **41**, 1953–1964.
8. Wan,L.C., Mao,D.Y., Neculai,D., Strecker,J., Chiovitti,D., Kurinov,I., Poda,G., Thevakumaran,N., Yuan,F., Szilard,R.K. *et al.* (2013) Reconstitution and characterization of eukaryotic N6-threonylcarbamoylation of tRNA using a minimal enzyme system. *Nucleic Acids Res.*, **41**, 6332–6346.
9. Perrochia,L., Guetta,D., Hecker,A., Forterre,P. and Basta,T. (2013) Functional assignment of KEOPS/EKC complex subunits in the biosynthesis of the universal t6A tRNA modification. *Nucleic Acids Res.*, **41**, 9484–9499.
10. Deutsch,C., El Yacoubi,B., de Crecy-Lagard,V. and Iwata-Reuyl,D. (2012) The biosynthesis of threonylcarbamoyl adenosine (t6A), a universal tRNA nucleoside. *J. Biol. Chem.*, **287**, 13666–13673.
11. Zhang,W., Collinet,B., Perrochia,L., Durand,D. and van Tilbeurgh,H. (2015) The ATP-mediated formation of the YgjD–YeaZ–YjeE complex is required for the biosynthesis of tRNA t6A in Escherichia coli. *Nucleic Acids Res.*, **43**, 1804–1817.
12. Hecker,A., Leulliot,N., Gadelle,D., Graille,M., Justome,A., Dorlet,P., Brochier,C., Quevillon-Cheruel,S., Le Cam,E., van Tilbeurgh,H. *et al.* (2007) An archaeal orthologue of the universal protein Kae1 is an iron metalloprotein which exhibits atypical DNA-binding properties and apurinic-endonuclease activity in vitro. *Nucleic Acids Res.*, **35**, 6042–6051.
13. Nichols,C.E., Lamb,H.K., Thompson,P., Omari,K.E., Lockyer,M., Charles,I., Hawkins,A.R. and Stammers,D.K. (2013) Crystal structure of the dimer of two essential Salmonella typhimurium proteins, YgjD & YeaZ and calorimetric evidence for the formation of a ternary YgjD-YeaZ-YjeE complex. *Protein Sci.*, **22**, 628–640.
14. Buss,K.A., Cooper,D.R., Ingram-Smith,C., Ferry,J.G., Sanders,D.A. and Hasson,M.S. (2001) Urkinase: structure of acetate kinase, a member of the ASKHA superfamily of phosphotransferases. *J. Bacteriol.*, **183**, 680–686.
15. Hecker,H., Graille,M., Madec,E, Gadelle,D., Lecam,E., van Tilbeurgh,H. and Forterre,P.2009)Theuniversal Kae1 protein and the associated Bud32 kinase (PRPK), a mysteriousprotein couple probably essential for genome maintenance in Archaea andEukarya. *Biochem Soc Trans.*, **37**, 29–35.
16. Hecker,A., Lopreiato,R., Graille,M., Collinet,B., Forterre,P., Libri,D. and van Tilbeurgh,H. (2008) Structure of the archaeal Kae1/Bud32 fusion protein MJ1130: a model for the eukaryotic EKC/KEOPS subcomplex. *EMBO J.*, **27**, 2340–2351.
17. Mao,D.Y., Neculai,D., Downey,M., Orlicky,S., Haffani,Y.Z., Ceccarelli,D.F., Ho,J.S., Szilard,R.K., Zhang,W., Ho,C.S. *et al.* (2008) Atomic structure of the KEOPS complex: an ancient protein kinase-containing molecular machine. *Mol. Cell*, **32**, 259–275.
18. LaRonde-LeBlanc,N. and Wlodawer,A. (2005) A family portrait of the RIO kinases. *J. Biol. Chem.*, **280**, 37297–37300.
19. Downey,M., Houlsworth,R., Maringele,L., Rollie,A., Brehme,M., Galicia,S., Guillard,S., Partington,M., Zubko,M.K., Krogan,N.J. *et al.* (2006) A genome-wide screen identifies the evolutionarily conserved KEOPS complex as a telomere regulator. *Cell*, **124**, 1155–1168.
20. Garay-Arroyo,A., Colmenero-Flores,J.M., Garciarrubio,A. and Covarrubias,A.A. (2000) Highly hydrophilic proteins in prokaryotes and eukaryotes are common during conditions of water deficit. *J. Biol. Chem.*, **275**, 5668–5674.
21. Corbacho,I., Olivero,I. and Hernandez,L.M. (2005) A genome-wide screen for Saccharomyces cerevisiae nonessential genes involved in mannosyl phosphate transfer to mannoprotein-linked oligosaccharides. *Fungal Genet. Biol.*, **42**, 773–790.
22. Pennock,E., Buckley,K. and Lundblad,V. (2001) Cdc13 delivers separate complexes to the telomere for end protection and replication. *Cell*, **104**, 387–396.
23. Martin,V., Du,L.L., Rozenzhak,S. and Russell,P. (2007) Protection of telomeres by a conserved Stn1-Ten1 complex. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 14038–14043.
24. Kisseleva-Romanova,E., Lopreiato,R., Baudin-Baillieu,A., Rousselle,J.C., Ilan,L., Hofmann,K., Namane,A., Mann,C. and Libri,D. (2006) Yeast homolog of a cancer-testis antigen defines a new transcription complex. *EMBO J.*, **25**, 3576–3585.
25. Juneau,K., Palm,C., Miranda,M. and Davis,R.W. (2007) High-density yeast-tiling array reveals previously undiscovered introns and extensive regulation of meiotic splicing. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 1522–1527.
26. Collinet,B., Friberg,A., Brooks,M.A., van den Elzen,T., Henriot,V., Dziembowski,A., Graille,M., Durand,D., Leulliot,N., Saint Andre,C. *et al.* (2011) Strategies for the structural analysis of multi-protein complexes: lessons from the 3D-Repertoire project. *J. Struct. Biol.*, **175**, 147–158.
27. Kabsch,W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.*, **26**, 795–800.
28. Schneider,T.R. and Sheldrick,G.M. (2002) Substructure solution with SHELXD. *Acta Crystallogr. D Biol. Crystallogr.*, **58**, 1772–1779.
29. Bricogne,G., Vonrhein,C., Flensburg,C., Schiltz,M. and Paciorek,W. (2003) Generation, representation and flow of phase information in

structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr. D Biol. Crystallogr.*, **59**, 2023–2030.
30. Murshudov,G.N., Vagin,A.A. and Dodson,E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.*, **53**, 240–255.
31. Perrakis,A., Morris,R. and Lamzin,V.S. (1999) Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.*, **6**, 458–463.
32. Adams,P.D., Grosse-Kunstleve,R.W., Hung,L.W., Ioerger,T.R., McCoy,A.J., Moriarty,N.W., Read,R.J., Sacchettini,J.C., Sauter,N.K. and Terwilliger,T.C. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.*, **58**, 1948–1954.
33. Vagin,A. and Teplyakov,A. (2000) An approach to multi-copy search in molecular replacement. *Acta Crystallogr. D Biol. Crystallogr.*, **56**, 1622–1624.
34. Terwilliger,T.C., Grosse-Kunstleve,R.W., Afonine,P.V., Moriarty,N.W., Zwart,P.H., Hung,L.W., Read,R.J. and Adams,P.D. (2008) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr. D Biol. Crystallogr.*, **64**, 61–69.
35. Murshudov,G.N., Skubak,P., Lebedev,A.A., Pannu,N.S., Steiner,R.A., Nicholls,R.A., Winn,M.D., Long,F. and Vagin,A.A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.*, **67**, 355–367.
36. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 2126–2132.
37. Kabsch,W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 125–132.
38. Sheldrick,G.M. (2010) Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 479–485.
39. McCoy,A.J., Grosse-Kunstleve,R.W., Adams,P.D., Winn,M.D., Storoni,L.C. and Read,R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.
40. Cowtan,K. (2010) Recent developments in classical density modification. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 470–478.
41. Winn,M.D., Ballard,C.C., Cowtan,K.D., Dodson,E.J., Emsley,P., Evans,P.R., Keegan,R.M., Krissinel,E.B., Leslie,A.G., McCoy,A. *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.*, **67**, 235–242.
42. Vagin,A. and Teplyakov,A. (2010) Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 22–25.
43. Konarev,P.V., Petoukhov,M.V., Volkov,V.V. and Svergun,D.I. (2006) ATSAS 2.1, a program package for small-angle scattering data analysis. *J. Appl. Crystallogr.*, **39**, 277–286.
44. Konarev,P.V., Volkov,V.V., Sokolova,A.V., Koch,M.H.J. and Svergun,D.I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.*, **36**, 1277–1282.
45. Kelley,L.A. and Sternberg,M.J. (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protoc.*, **4**, 363–371.
46. David,G. and Perez,J. (2009) Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline. *J. Appl. Crystallogr.*, **42**, 892–900.
47. Svergun,D.I., Petoukhov,M.V. and Koch,M.H. (2001) Determination of domain structure of proteins from X-ray solution scattering. *Biophys. J.*, **80**, 2946–2953.
48. Facchin,S., Lopreiato,R., Stocchetto,S., Arrigoni,G., Cesaro,L., Marin,O., Carignani,G. and Pinna,L.A. (2002) Structure-function analysis of yeast piD261/Bud32, an atypical protein kinase essential for normal cell life. *Biochem. J.*, **364**, 457–463.
49. Petoukhov,M.V. and Svergun,D.I. (2005) Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophys. J.*, **89**, 1237–1250.
50. Moore,B.L., Kelley,L.A., Barber,J., Murray,J.W. and MacDonald,J.T. (2013) High-quality protein backbone reconstruction from alpha carbons using Gaussian mixture models. *J. Comput. Chem.*, **34**, 1881–1889.
51. Svergun,D., Barberato,C. and Koch,M.H.J. (1995) CRYSOL—a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.*, **28**, 768–773.
52. Hecker,A., Leulliot,N., Gadelle,D., Graille,M., Justome,A., Dorlet,P., Brochier,C., Quevillon-Cheruel,S., Le Cam,E., van Tilbeurgh,H. *et al.* (2007) An archaeal orthologue of the universal protein Kae1 is an iron metalloprotein which exhibits atypical DNA-binding properties and apurinic-endonuclease activity in vitro. *Nucleic Acids Res.*, **35**, 6042–6051.
53. Schneidman-Duhovny,D., Inbar,Y., Nussinov,R. and Wolfson,H.J. (2005) PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res.*, **33**, W363–W367.
54. Miyoshi,A., Kito,K., Aramoto,T., Abe,Y., Kobayashi,N. and Ueda,N. (2003) Identification of CGI-121, a novel PRPK (p53-related protein kinase)-binding protein. *Biochem. Biophys. Res. Commun.*, **303**, 399–405.
55. LaRonde-LeBlanc,N. and Wlodawer,A. (2005) The RIO kinases: an atypical protein kinase family required for ribosome biogenesis and cell cycle progression. *Biochim. Biophys. Acta*, **1754**, 14–24.
56. Ferreira-Cerca,S., Sagar,V., Schafer,T., Diop,M., Wesseling,A.M., Lu,H., Chai,E., Hurt,E. and LaRonde-LeBlanc,N. (2012) ATPase-dependent role of the atypical kinase Rio2 on the evolving pre-40S ribosomal subunit. *Nat. Struct. Mol. Biol.*, **19**, 1316–1323.
57. Nilsson,J. and Nissen,P. (2005) Elongation factors on the ribosome. *Curr. Opin. Struct. Biol.*, **15**, 349–354.
58. Costessi,A., Mahrour,N., Sharma,V., Stunnenberg,R., Stoel,M.A., Tijchon,E., Conaway,J.W., Conaway,R.C. and Stunnenberg,H.G. (2012) The human EKC/KEOPS complex is recruited to Cullin2 ubiquitin ligases by the human tumour antigen PRAME. *PLoS One*, **7**, e42822.